Data Analysis Plan

Study Protocol

Phase II clinical trial testing the safety and efficacy of 3,4-methylenedioxymethamphetamine (MDMA)-assisted psychotherapy in subjects with chronic posttraumatic stress disorder.

Protocol Amendment #5: January 23, 2009

Data Analysis Plan Version 1: March 16, 2011
Data Analysis Plan Version 2: May 3, 2011
Data Analysis Plan Version 3: November 9, 2011

Sponsor:

Multidisciplinary Association for Psychedelic Studies (MAPS)
309 Cedar Street # 2323
Santa Cruz, CA 95060

Principal Investigator:

Michael C. Mithoefer M.D.

Co-Investigators:

Mark T. Wagner Ph.D.
Ann T. Mithoefer, BSN

Prepared by:

L. (Ilsa) Jerome, Ph.D.

Table of Contents

Data Analysis Plan for MP-1

1.0 Definitions

Categorical data: This refers to discrete (indivisible) variables, such as gender or ethnicity. These data will be presented as total numbers of each category as needed to describe the sample.
Descriptive data: This includes mean, median, standard deviation, minimum and maximum of numerical data used as needed to describe the sample.
Difference scores: These consist of scores computed by subtracting one value from another, as subtracting baseline from End of Stage 1 score, used to test for differences between and within groups to determine change as a function of experimental treatment over time.
Efficacy: A type of analysis used to assess therapeutic effects or benefits.
Exploratory analyses: An inferential or descriptive analysis of the data to determine trends that might lead to hypotheses for further study.
Frequency listing: A tabular listing of numbers and/or percentages of events used as needed to describe the sample or data characteristics.
Outcome measures: These are primary and secondary study measures that are used to test the study hypotheses.
Process measures: These are study measures or qualitative observations collected during the study that may increase depth of understanding and that are not necessarily related to safety or efficacy.
Protocol deviation: An event that represents significant divergence from the intended study design as described in the protocol.
Safety: An assessment of the condition of study subjects that examines potential risks, adverse events and reactions.
Safety measures: These are study measures that assess safety, such as blood pressure monitoring. These measures are used to assess safety of the study drug.
Spontaneously reported reactions, reactions: Specific expected reactions gathered from the literature on MDMA, referred to as side effects in the study protocol.
Study design: All elements of a research project that define the study question, experimental methods, study procedures including blinding and randomization, measurement techniques, flow sheet of data, and statistical analysis.
Tabular Listing: A list of each variable or item for each individual subject either in total or by condition in a table format.

2.0 Introduction

This is a data analysis plan for the study “Phase II clinical trial testing the safety and efficacy of 3,4-methylenedioxymethamphetamine (MDMA)-assisted psychotherapy in subjects with chronic posttraumatic stress disorder (PTSD).”

3.0 Study Objectives

3.1 Primary Objective

Volunteers receiving MDMA-assisted psychotherapy will experience (trends toward) a greater decrease in signs and symptoms of PTSD than controls at two months after the second drug-assisted (MDMA or placebo) session. The primary outcome measure evaluating efficacy will be the Clinician-Administered PTSD Scale (CAPS).

- To evaluate changes in PTSD symptoms as measured via CAPS scores at baseline before treatment and again two months after the second experimental session in subjects receiving the placebo vs. full dose of MDMA-assisted psychotherapy.

3.2 Secondary Objectives

Volunteers receiving MDMA-assisted psychotherapy will experience (trends toward) a greater decrease in signs and symptoms of PTSD than controls after each experimental session, as measured by the Clinician-Administered PTSD Scale (CAPS), the self-reported Impact of Events Scale (IES-R) and Symptoms Checklist-90-R (SCL90-R). Secondary outcome measures evaluating efficacy will be the IES-R and the SCL90-R.

- To evaluate changes in PTSD symptoms as measured via CAPS scores after each experimental session in subjects receiving the placebo vs. full dose of MDMA-assisted psychotherapy.
- To evaluate the extent to which a given stressful (traumatic) life event produces subjective distress via IES-R after each experimental session in subjects receiving the placebo vs. full dose of MDMA-assisted psychotherapy.
- To evaluate psychological symptoms and affective states via the SCL90-R after each experimental session in subjects receiving the placebo vs. full dose of MDMA-assisted psychotherapy.

3.3 Safety Objectives

MDMA would be well tolerated by people with PTSD receiving the study drug in combination with psychotherapy.

Exposure to MDMA will not be associated with neurocognitive toxicity as assessed by the Repeatable Battery for Assessment of Neuropsychological Status (RBANS), the Paced Auditory Serial Addition Task (PASAT) and the Rey-Osterrieth Complex Figure Test (Rey CFT). The assessment of neuropsychological status serves as a means of safety evaluation. The NEO Personality Inventory is a neuropsychological measure of personality that is also a part of our safety evaluation.

- To assess adverse events throughout the study.
- To assess spontaneously reported reactions (“side effects” in protocol) on during and 7 days after each experimental session.
- To assess blood pressure and pulse during experimental sessions using automated blood pressure and pulse monitoring equipment.
- To assess body temperature at regular intervals during experimental sessions.
- To assess experience of degree of psychological distress by repeated administration of the Subjective Units of Distress (SUD) during experimental sessions.

- To detect any changes in liver function at baseline after the second experimental session or two months after the second experimental session by assays of liver enzymes in the blood.
- To evaluate neurocognitive function at baseline and two months after the second experimental session via the RBANS.
- To evaluate neurocognitive function at baseline and two months after the second experimental session via the PASAT.
- To evaluate neurocognitive function at baseline and two months after the second experimental session via the Rey CFT.
- To evaluate changes in personality traits comparing baseline and two months after the second experimental session via the NEO.

4.0 Study design

The study followed a randomized, double-blind placebo controlled design, with psychotherapists and independent raters blinded to participant condition. Subjects were assigned to receive either psychotherapy with MDMA or lactose placebo. Study amendments later included the addition of a supplemental dose that could be given 2 to 2.5 hours after the initial dose, the addition of an open-label study segment (Stage 2) for subjects assigned to receive placebo, the addition of an open label session to occur after the blind was broken for subjects assigned to the MDMA condition, the enrollment of a veteran who had not received psychotherapy or pharmacotherapy, and the addition of a long-term follow up conducted at least 10 months after completion of Stage 1 or Stage 2, as appropriate.

Planned enrollment was for twelve subjects in the MDMA condition and eight subjects in the placebo condition, with dropouts replaced until 20 subjects had completed the study. The amended study also enrolled an additional subject who was not treatment-resistant, also randomized to one of the two conditions.

4.1 Time and Events table

Please see attached document.

5.0 Measures

5.1 Outcome Measures

Clinician-Administered PTSD Scale (CAPS) Global Score, subscale scores (B, C, D, F), Diagnostic criteria Met score, Associated Features (#26, 27, 28, 29, 30)
Impact of Events Scale-Revised (IES-R), Total score
NEO Personality Inventory, Trait scores (Neuroticism, Extroversion, Openness, Conscientiousness, and Agreeableness)
Symptom Checklist 90-Revised (SCL90-R) GSI, PST, PSDI
Long-term Follow-up Questionnaire Benefits (Questions 1-6), Current Psychotherapy (Question 14), Current medications (Question 15)

5.2 Safety Measures

Paced Auditory Serial Addition Task (PASAT) Trial 1 score, Trial 2 Score
Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) Global and scale scores (Memory, Visuospatial, Language, Attention, Delayed Recall)
Rey-Osterrieth Complex Figure Test (RCFT) 30 second delay
Standard assay of liver enzymes
Subjective Units of Distress (SUD)
Vital signs (Systolic blood pressure (SBP), diastolic blood pressure (DBP), heart rate (HR), body temperature (BT))
Long-term Follow-up Questionnaire Harms (Questions 7-12) Memory effects (Question 18)

Spontaneously reported reactions during experimental sessions and seven days after
Adverse events reported during the course of the study
General well-being assessment

5.3 Process Measures

Reactions to Research Participation Questionnaire – Short Form Revised (RRPQ)
Subject Belief of Condition Assignment
Working Alliance Inventory (WAI)

6.0 Analyses

In general, nominal variables will be described in terms of frequencies and percentages and analyzed using chi square analysis. Ordinal and non-normal continuous variables will be described using sample median and range, and analyzed by non-parametric statistical tests, and approximately normal variables will be described using sample mean and standard deviations and analyzed by parametric statistical tests. Appropriate tests for comparative group homogeneity would be conducted and any significant lack of homogeneity will be appropriately addressed

6.1 Study population

See protocol. All subjects were diagnosed with chronic PTSD, met DSM-IV criteria, and had Global CAPS scores of 50 or greater upon enrollment, and 20 of 21 had undergone psychotherapy and pharmacotherapy without symptom reduction. The index trauma could either be crime-related or combat-related.

All clinical data will be presented in tabular listings. All analyses will be carried out with SPSS of Version 12.0 or higher.

Definitions of subject populations for analysis:

All Enrolled: All subjects who signed an informed consent form and completed baseline measures.

Intention to treat: All subjects who were randomized to a condition and underwent at least one experimental session. All available data will be used.
Per Protocol: All subjects who completed Stage 1 and underwent assessment of PTSD symptoms two months after the second experimental session. Analyses may be conducted with and without a participant who corresponds to a major protocol deviation.
Partial crossover: All subjects who completed Stage 2 or an open-label MDMA session in addition to completing Stage 1.

6.2 Protocol Deviations

All protocol deviations will be included as a categorized listing. Subjects with minor deviations will be included in all analyses. Analyses will be performed with and without deviations to examine the effects of including them in an analysis. Secondary analyses may be conducted to examine interactions with certain characteristics within the subject population. If it is appropriate as indicated via analyses, subjects with major deviations will be excluded from the per protocol analysis and included in the intention to treat analysis. Safety analyses will include all enrolled subjects with all available data.

Possible deviation categories include:

- Entered study but did not meet entry criteria
- Developed withdrawal criteria during the study but were not withdrawn
- Received wrong treatment or incorrect dose
- Received excluded concomitant treatment
- Protocol procedure not performed per protocol
- Protocol procedure performed out of range
- Informed consent performed not per protocol

6.3 Participant Demographics and Background

Population: All enrolled and Per protocol

Categorical Data includes: Gender, ethnicity/race, trauma etiology, whether the subject was a patient of the PI, medical history, psychiatric history, physical examination, lab values, general well being

Descriptive Data includes: Age, Number of years with PTSD, number and duration of past therapy for PTSD, number and duration of past medications for PTSD, number of incidences of prior ecstasy use

Format of presentation: Summary tables including frequency listings in total and by condition

Content of presentation: Gender, age, ethnicity/race, trauma etiology, number of years with PTSD, number and duration of past therapy for PTSD, number and duration of past medications for PTSD, number of incidences of prior ecstasy use, whether the subject was a patient of the PI, percent co-morbidity (as computed by summing number of current psychiatric diagnoses other than PTSD.)

Demographic factors analysis

Goal: To test for a main effect of the basic demographic factors gender, age, number of years of psychotherapy and presence of additional affective disorders upon PTSD symptoms across both conditions, meeting PTSD diagnostic criteria, and reduced psychological symptoms at the end of Stage 1

P*opulation*: Intention to treat, per protocol

Data included in analysis: CAPS (Global Score and subscales, Diagnostic criteria met, Associated Features), IES-R Global score, SCL90-R Global scores (GSI, PSDI, PST). Demographic variables will include gender, age, years of psychotherapy, presence of other affective disorder.

Format of analysis: Repeated-measures ANOVA for categorical data (gender, presence of additional affective disorder), linear regression for continuous data (age, years of psychotherapy)

Time: Baseline, End of Stage 1

Between group factors / independent predictors: Gender, age, years of psychotherapy, presence/absence of other affective disorder

6.4 Efficacy Analyses

Descriptive Data includes: CAPS (Global Score, selected subscale scores, Diagnostic criteria met, Associated Features), IES-R (Global Score), SCL90-R (Global Score and Subscales)

Format of presentation: Summary tables including descriptive data presented in total and by condition. Descriptive data will also be provided for any additional analyses. When applicable, summary tables divided on the basis of a given demographic variable will be provided, when demographic variables may serve as the basis for additional or subsidiary analyses, as participant age, gender or status as a patient of the PI.

Content of presentation: CAPS (Global Score and subscales, Diagnostic criteria met, Associated Features), IES-R (Global Score), SCL90-R (Global Score and Subscales). Analyses will only be presented in text if a notable interaction is found.

6.4.1 Main analyses

Goal: To determine if there is a main effect of condition on PTSD symptoms, criteria for meeting PTSD, and psychological symptoms over the course of Stage 1 (after each experimental session) and at the end of Stage 1

Population: Per protocol

Data included in analyses: CAPS (Global Score, subscales, Diagnostic criteria met), IES-R Global score, SCL90-R Global scores (GSI, PSDI, PST)

Format of analyses: Repeated-measures ANOVA. Bonferroni corrections will be applied if multiple scores measuring the same parameter from a single scale are assessed.

Time: Baseline, after Exp. Session 1, after Exp. Session 2, End of Stage 1

Between-groups factors: Condition (MDMA, placebo)

6.4.2 Additional analyses

Analysis #1

Goal: To determine if there is a main effect of condition and participation in partial crossover upon PTSD symptoms, meeting PTSD diagnostic criteria, and reduced psychological symptoms at the end of Stage 1 and at the end of the partial crossover

Population: Partial Crossover

Data included in analyses: CAPS (Global Score and subscales specified above, Diagnostic criteria met, Associated Features), IES-R Global score, SCL90-R Global scores (GSI, PSDI, PST) – only if significant findings in main analyses

Format of Analyses: Repeated measures ANOVA. Bonferroni corrections will be applied if multiple scores measuring the same parameter from a single scale are assessed.

Time: Baseline, End of Stage 1 or re-baseline assessment when applicable for Stage 2 subjects, End of Stage 2

Between-groups factors: Condition (MDMA, placebo)

Analysis #2

Goal: To test for an effect of condition and presence of a third open-label session on PTSD symptoms, meeting PTSD diagnostic criteria, and psychological symptom at the end of Stage 1 and after third open-label session. This analysis will only be performed upon scores that showed significant differences between baseline and end of stage 1.

Population: , Per Protocol (restricted to subjects with third open label session) and Partial Crossover

Data included in analyses: CAPS (Global Score and subscales specified, Diagnostic criteria met, Associated Features), IES-R Global score, SCL90-R Global scores (GSI, PSDI, PST) – only if significant findings in main analyses

Format of Analyses: Repeated measures ANOVA. Bonferroni corrections will be applied if multiple scores measuring the same parameter from a single scale are assessed.

Time: Baseline, End of Stage 1, after Third Open Label Session

Between-groups factors: Condition (MDMA, placebo)

Analysis #3

Goal: To test the hypotheses that medication tapering has an effect on PTSD symptoms as assessed via CAPS. Will be performed only if there are significant effects of time or condition on a given CAPS score.

Population: Per Protocol

Data included in analysis: Baseline (pre and post tapering) CAPS Global Scores, to detect potential variation in PTSD symptoms due to psychotherapy and medication tapering

Format of Analysis: chi square

Time: Baseline pre-tapering (Visit 1), Baseline post-tapering (Visit 4)

Between-groups factor: Tapering (With tapering, without tapering)

6.4.3 Subsidiary analyses

Goal: To test for the source of variance of any potential interactions between demographic variables and outcome measures, and to test whether there is an effect from being a patient of the PI prior to study participation, including any interactions between this specific demographic variable and study condition. With the exception of analyses relating to subjects who were and were not patients of the PI, these analyses will only be performed if any interactions are found between specific criteria and efficacy data. NOTE: if there are no significant findings for given score, will not perform subsidiary analysis on specific score.

Population: and Per Protocol

Data included in analyses: CAPS (Global Score and specified subscales, Diagnostic criteria met, Associated Features), IES-R Global score, SCL90-R Global scores (GSI, PSDI, PST)
If and only if there are systematic differences detected by chi square in a demographic variable (age, gender, # years psychotherapy), patient or not patient of PI.

Format of Analyses: A) Repeated measures analysis of variance using condition and demographic variable as between group variable and time of administration as repeated measure OR B) two separate repeated measures analyses of variance with condition as a between-group factor, one analyses with and one without members of specific category (such as treatment-

resistance). Comparisons of patients of the PI and non-patients will be performed as two separate analyses that include and exclude patients of the PI.

Time: Baseline, After Experimental Session 1 (Stage 1), After Experimental Session 2 (Stage 1), End of Stage 1

Between-groups factors: Condition (MDMA, placebo); Demographic factor (e.g. male, female) Previous patient of the PI (Yes/No)

6.4.4 Long-term Follow Up Analyses

Categorical Data includes: Long-term follow-up Questionnaire items

Descriptive Data includes: CAPS (Global Score and selected subscale scores, Diagnostic Criteria met, Associated Features), IES-R

Format of presentation: Summary tables with frequency listings in total and by condition. Summary tables of descriptive data in total and by condition.

Content of presentation:
Categorical data, including: Benefits, including Frequencies of each type of benefit reported Frequencies of any benefit (Y/N), degree of benefit, benefits lasting to present, how much benefit lasted;
Harms, including Frequencies of any harms, degree of harm, harms lasting to present, how much harms lasted; Frequencies of subjects reporting belief that additional sessions would be helpful and time at which additional sessions would be helpful (soon after, later); Frequency of subjects taking medication (type of medication, self-reported reason for prescription, taking at baseline, taking during follow-up period, taking at time of LTFU); Frequency of subjects in therapy (type of therapy, therapy at baseline, during follow-up period, same or different as baseline) Analyses will only be presented in text if a notable interaction is found.

<u>Main analysis:</u>

Goal: To test if MDMA-assisted psychotherapy continues to have a main effect upon PTSD symptoms at least ten months later

Population: Per Protocol

Data included in analyses: CAPS (Global Score and specified subscales, Diagnostic criteria met, Associated Features), IES-R Global Score

Format of Analysis: Repeated measures ANOVA using time as repeated measure

Time: Baseline, End of Stage 1, Long-term follow-up

Between-groups factors: None, combining data from MDMA and Placebo conditions since all but one of the subjects received MDMA, either in Stage 1 or Stage 2.

Subsidiary Analyses:

Will be conducted if and only if there are a) systematic differences detected by chi square or b) in a demographic variable or c) could provide essential information on effects of study drug.

Analysis #1

Goal: To test if there is a main effect of condition on PTSD symptoms at long-term follow-up

Population: Per Protocol

Data included in analyses: CAPS (Global and Subscale Scores, Diagnostic criteria met, Associated Features), IES-R Global Scores

Format of Analyses: One-way ANOVA on LTFU scores with Condition (MDMA, placebo) as between-groups factor.

Time: Baseline, End of Stage 1, Long-term Follow Up

Between-groups factor: Condition (MDMA, placebo)

Analysis #2

Goal: To test the presence of an effect of time upon PTSD symptoms that remains at least 12 months after study completion. These analyses will be performed if there is no effect of condition at follow-up in Analysis #1

Population: Per Protocol

Data included in analyses: CAPS (Global Score and specified subscale scores, Diagnostic criteria met, Associated Features), IES-R Global Score

Format of Analyses: Repeated measures ANOVA using time of administration as repeated measure and pooling across initially assigned conditions.

Time: Baseline, End of Stage 1, Long-term Follow Up

Between-groups factor: Condition assigned at Stage 1 (MDMA or Placebo); Demographic factors (gender, age, # of years of psychotherapy or pharmacotherapy, patient or not patient of PI)

Exploratory analyses

Analysis #1

Goal: To test for the presence of interactions between demographic variables and outcome measures and self-reported benefits at long-term follow up. If outcome or safety analyses uncover significant impact from demographic variables, especially in interactions. If there are no significant differences due to condition, then data will be pooled across conditions.

Population: Per Protocol

Data included in analyses: Long-term Follow-up Questionnaire, CAPS (Global Score and subscales, Diagnostic Criteria met, Associated Features), IES-R Global Score

Format of analysis: One-way or two-way ANOVA or correlational analysis (as appropriate)

Between-group factors: Condition (MDMA, Placebo); Demographic variables (gender, age, # of years of psychotherapy or pharmacotherapy, patient or not patient of PI)

Analysis #2

Goal: To examine PTSD diagnosis using the "intention to treat" sample

Population: Intention to treat

Data included in analysis: CAPS diagnostic score at end of Stage 1

Format of analysis: Chi-square

Between-group factors: Condition (MDMA, Placebo)

6.5 Safety Analysis

Population: All enrolled, intention to treat, per protocol, partial crossover

Categorical data includes: Spontaneously reported reactions occurring seven days after an experimental session (maximum intensity, duration within 24 hrs of the experimental session, total duration across seven days), Adverse Events

Format of presentation: Summary tables of frequency listings in total, by condition (placebo, MDMA, partial crossover), by relatedness to study drug, severity (mild, moderate, severe) by session number (1st, second, third when applicable), by body system, by classification

Content of presentation: Specific spontaneously reported reactions by classification (across 7 days after an experimental session), number of spontaneously reported reactions occurring on seven days, Adverse Events

Descriptive data includes: PASAT (Raw and Percentile scores for Trial 1 and Trial 2, change scores); RBANS (Total Raw and Percentage scores, raw subscale scores); Rey CFT 30-second delay (Raw, Time, Percentage scores); NEO personality trait scales (Openness,

Conscientiousness, Extroversion, Agreeableness, Neuroticism scores); Physiological data (vitals, vital signs above cut-off and time point above cut-off during experimental sessions); psychological distress (SUD); Lab values; General Wellbeing

Format of presentation: Summary tables of frequency listings in total, by condition and by session number when applicable.

Content of presentation: PASAT (Raw and Percentage scores for Trial 1 and Trial 2, change scores); RBANS (Total Raw and Percentage scores, raw subscale scores); Rey CFT 30-second delay (Raw, Time, Percentage scores); NEO personality trait scales (Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism scores); Pre-drug average, maximum change (peak), and post-drug average values of physiological data (HR, SBP, DBP, BT); vital signs above cut-off and time point above cut-off during experimental sessions; psychological distress (SUD); Liver panel values (Baseline, four to seven days after the second experimental session and/or the two months after the second experimental session); General Wellbeing (descriptives for each day by condition)

6.5.1 Main Analysis

Physiological Measures and SUD

Analysis #1

Goal: to test the presence of a main effect of condition (MDMA, placebo) on changes in vital signs and psychological distress,

Population: All enrolled, intention to treat, per protocol, partial crossover

Data included: Pre-drug average, maximum change (peak) and post-drug average SBP, DBP, HR, BT and SUD, number of subjects with vital signs above clinical cut off.

Format of Analyses: One-way ANOVA per each experimental or partial crossover (Stage 2 or third open label) session

Between-Group Factor: Condition (MDMA, placebo)

Analysis #2

Goal: To test effects of condition, presence or absence of supplemental dose and time of session on physiological data and psychological distress during experimental sessions

Population: All enrolled, intention to treat, per protocol

Data included: Pre-drug average, maximum change (peak) and post-drug average SBP, DBP, HR, BT and SUD, number of subjects with vital signs above clinical cut off.

Format of analysis: Repeated measures ANOVA

Time: Session (Experimental 1, Experimental 2)

Between Groups Factors: Condition (MDMA, placebo) as between-group factor, and supplemental dose (given, not given) as between-group factor.

Analysis #3

Goal: To test effects of condition, presence or absence of supplemental dose and time of session on physiological data and psychological distress during partial crossover (Stage 2 and third open label) sessions.

Population: Partial crossover

Data included: Pre-drug average, maximum change and post-drug average SBP, DBP, HR, BT and SUD, number of subjects with vital signs above clinical cut off.

Format of analysis: One-way ANOVA

Time: Crossover Session (Stage 2 Session 1, Stage 2 Session 2, post-Stage 1Third Open Label Session)

Between-Subjects factors: Supplement (present, absent)

Neurocognitive Function:

Goal: To determine if MDMA has a detrimental effect upon neurocognitive function; to test if there is a main effect of condition upon neurocognitive function measure scores

Population: Per protocol

Descriptive data included: PASAT trial 1 and trial 2 Percentage scores, RBANS total scores, Rey-Osterreith 30 second recall scores

Format of analyses: Independent sample t-test comparing Baseline scores.
If no differences are detected between scores at Baseline, then a second independent sample t-test comparing neurocognitive function scores at End of Stage 1.
If there are significant differences at Baseline, then a second independent sample t-test will be performed upon difference scores computed for the test scores exhibiting these differences.
NOTE: Difference scores will be computed prior to this analysis

Between-group variables: Condition (MDMA, placebo)

Laboratory Values:

Goal: To test for an effect of condition (MDMA or placebo) upon liver panel scores

Population: Intention to treat, Per protocol

Data included: All value reported in the standard liver panel, including ASG, ALT, BUN, bilirubin, and BUN/creatine ratio.

Format of analysis: Initial independent sample t-test comparing each liver panel score at baseline, with Bonferroni correction used to address multiple scores from same measure.
If no significant differences are found, initial tests will be followed by subsequent independent sample t-tests at Visit 13, with Bonferroni correction used to address multiple scores from same measure when applicable.
If there are significant between-group differences one or more between-group laboratory value, then difference scores will be computed for that value and a subsequent independent sample t-test will be performed upon the difference scores, with Bonferroni correction used to address multiple scores from same measure.

Between-group factor: Condition (MDMA, placebo)

NEO Scores:

Analysis #1

Goal: To test for an effect of MDMA upon facets of personality, as assessed via all five NEO personality scores (Neuroticism, Extroversion, Conscientiousness and Agreeableness)
Population: Intention to treat, per protocol

Data included: NEO scale scores for Baseline and at the end of Stage 1.

Format of analysis: Initial independent sample t-test comparing each NEO scale score at baseline
If no significant differences are found, initial tests will be followed by subsequent independent sample t-tests at two months after the second experimental session.
Between-groups factor: Condition (MDMA, placebo)

Analysis #2

Goal: To test for an effect of MDMA upon NEO scale scores as assessed by comparing NEO scores at baseline, at the end of Stage 1, and at the end of Stage 2

Population: Per protocol, partial crossover

Data included: All NEO scale scores at baseline, at the end of Stage 1 and at the end of Stage 2

Format of Analysis: Repeated measures ANOVA

Time: Baseline, End of Stage 1, End of Stage 2

Between-group factor: Condition (MDMA, placebo)

6.5.2 Subsidiary analyses

Physiological and SUD:

Analysis #1

Goal: To test for the presence of interactions between demographic variables and condition on one or more physiological variable or psychological distress. Performed if and only if there are significant differences in representation of a given demographic variable or feature across conditions (as, more women in one condition than the other), or if doing so gathers essential information on study drug safety.

Population: All enrolled, intention to treat, per protocol, partial crossover

Data included in analyses: Peak SBP, DBP and HR, peak SUD

Format of Analysis: Two-way analysis of variance, one per Stage 1 experimental sessions, for one partial crossover (Stage 2) sessions

Between-group factors: Demographic variable (e.g. gender), Condition (MDMA/placebo).

Analysis #2

Goal: To determine if there is an effect for Stage (e.g. order of administration) upon measures of physiological response or psychological distress taken during full-dose MDMA sessions.

Population: Intention to treat, per protocol [MDMA condition only], partial crossover

Data included in analyses: Pre-drug average, peak and post-drug average vital signs computed across all subjects averaged across 1) Randomized or 2) Open label full dose sessions

Format of analysis: One way between-group ANOVA

Between group factor: (Randomized MDMA subjects, Open Label)

Neurocognitive Function:

Goal: To test whether condition (MDMA, placebo) affects performance on individual RBANS subscales

Population: per protocol

Data included: RBANS subscale scores (Memory, Visuospatial, Language, Attention, Delayed Memory)

Format of analyses: Independent sample t-test comparing Baseline scores.
If no differences are detected between scores at Baseline, then a second independent sample t-test comparing neurocognitive function scores at End of Stage 1.
If there are significant differences between at Baseline, then a second independent sample t-test will be performed upon difference scores computed for the test scores exhibiting these differences. NOTE: Difference scores will be computed prior to this analysis

Between-group variables: Condition (MDMA, placebo)

6.5.3 Adverse Events

Summary Tables

Goal: To present adverse events organized by condition, seriousness, severity and relatedness to the study drug

Population: All enrolled, Intention to treat, Per protocol, Partial crossover

Categorical data includes: Number of AEs per subject; Number of severe AEs; Instances of occurrence of each spontaneously reported expected adverse event by dose received (placebo or MDMA); Adverse events (AE identity, seriousness (Yes/No), severity (mild, moderate, severe), onset, resolution, relatedness to study drug (not related, possibly related, probably related)); Spontaneously reported reactions listed by MedDRA System Order Class

Format of presentation: This data will be presented in several summary tables of counts per subject, by condition or across the course of the study, as appropriate for the data or descriptive statistics being listed.

Content of Presentation: Number of AEs per subject; Number of severe AEs; Instances of occurrence of each spontaneously reported expected adverse event by condition (MDMA, Placebo) and Stage (Stage 1, Stage 2); Adverse events (AE identity, seriousness (Yes/No), severity (mild, moderate, severe), onset, resolution, outcome, relatedness to study drug (not related, possibly related, probably related); Spontaneously reported reactions listed by MedDRA System Order Class

Descriptive data includes: Overall number of AEs, severity of AEs, Peak Demeanor after each experimental and open label session, average reported severity and duration of spontaneously reported reactions

Format of Presentation: Summary tables of frequency listings in total, by condition, by session, by subject and by condition

Content of Presentation: Adverse Event Listings (All possibly and probably related AEs, all AEs occurring within seven days of drug administration, all severe AEs, and number of severe AEs by condition); AE or severity of AE and peak demeanor after each experimental or open label session by condition (MDMA, placebo, partial crossover).

Tabular Listing of Adverse Events

Goal: To follow guidance for display of adverse events

Population: All enrolled, intention to treat, per protocol, partial crossover

Data included: Per each AE; Subject number, Patient identifier
- Age, race, sex, weight (height, if relevant)
- Location of CRFs, if provided
- Adverse event description (preferred term, reported term)
- Duration of the adverse event
- Severity (e.g., mild, moderate, severe)
- Seriousness (serious/non-serious)
- Action taken (none, dose reduced, treatment stopped, specific treatment instituted etc.)
- Outcome (e.g., CIOMS format)
- Relatedness process of determination given if possible
- Date of onset or date of clinic visit at which the event was discovered
- Timing of onset of the adverse event in relation to last dose of investigational product (if or when applicable)
- Investigational product dose in absolute amount, mg/kg
- Duration of investigational product treatment (Stage 1, Stage 2, third open label session)
- Concomitant treatment during study.

Format of presentation: All adverse events for each patient, including the same event on several occasions will be listed, giving both preferred term and the original term used by the investigator. The listing should be by treatment group and will include:

Analysis #1

Goal: To see if one or more demographic factor and condition alter the number and severity of adverse events collected across the course of the study.

Population; All enrolled, intention to treat, per protocol, partial crossover

Data included: Number of AEs collected across Stage 1 and Stage 2, Number of severe AEs collected across Stage 1 and partial crossover (Stage 2 or third open label session)

Format of analysis: Two-way ANOVA for gender, correlation or regression analysis for age

Between-groups factors: Dose administered (Placebo, MDMA); gender; age

Analysis #2

*Goa*l: to see if there is an effect of presence versus absence of supplemental dose upon number of AEs and number of severe AEs

Population: All enrolled, intention to treat, per protocol, partial crossover

Data included: Number of AEs across all study stages, number of severe AEs across all study stages

Format of analysis: One-way ANOVA

Between groups factor: Dosage (placebo, 125 mg, 187.5 mg)

Analysis #3

Goal: To see if demographic variables interact with dose of study drug; performed if and only if a previous analysis establishes that gender, age or presence of another affective disorder produce an interaction with condition.

Population; All enrolled, intention to treat, per protocol, partial crossover

Data included: Number of AEs across all study stages, number of severe AEs across all study stages.

Format of analysis: Two-way ANOVA for categorical data, regression analysis for continuous data

Between-group factors: Demographic factor (gender (male, female), age, presence of another affective disorder), study drug dose (placebo, 125 mg, 187.5 mg)

6.5.4 Long-Term Follow Up Questionnaire

Data includes: Any harms, (Y/N), degree of harms, Duration (harms lasted), Number of harms subject listed, type of harms listed, Taken ecstasy (Y/N), number of other drugs (0-n, where n = all listed), Perceived changes / lack of changes in cognitive function, NEO scales (Neuroticism, Extroversion, Openness, Conscientiousness, and Agreeableness)

Format of presentation: Summary tables of frequency listings in total and by condition for each long-term questionnaire variable

Content of Presentation: Each type of harm reported (as individually indicated in questionnaire), Any harm, degree of harm, harms lasting to present, how much harm.

Exploratory Analysis

Goal: To test for the presence of interactions between demographic variables and measures of self-reported harms at long-term follow up. To test if outcome or safety analyses uncover significant impact from demographic variables, especially in interactions. If there are no significant differences due to condition, then data will be pooled across conditions

Population: Per protocol

Format of analysis: One-way or two-way ANOVA or correlational analysis (as appropriate)

Between-group factors: Condition (MDMA, Placebo); Demographic variables (gender, age, # of years of psychotherapy or pharmacotherapy, patient or not patient of PI)

7.0 Process / Non-Outcome Measures

Descriptive listings will be presented in total and by condition for the data listed below.

Data includes: RRPQ Scores (Positive, Negative effects), WAI Scores (Factor 1, Factor 2, Factor 3, Global), In Therapy Now, Therapy Same/Different.

Analyses

Goal: To determine the strength and degree of the study blind

P*opulation*: Per protocol and Intention to treat

Format of Analysis: chi-square or other categorical analysis or informal analyses if data precludes these analyses

Between-group factors: Condition (MDMA, placebo); guesses of condition assignments by participant and therapists after each experimental session; guesses made at the end of Stage 1 by subjects and investigators.

8.0 Interim Analyses

Blinded participant data was examined but not analyzed during five scheduled Data Safety Board meetings.

Preliminary data spanning outcome measures and reactions occurring from study enrollment to the end of Stage 2 were analyzed and published in 2010. No preliminary analysis has been performed on long-term follow up data.